

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

On protein - protein search, using sw model

Run on: January 27, 2005, 19:22:00 ; Search time 191 seconds
(without alignmentB)

33.137 Million cell updates/sec

Title: US-09-830-837a-77
Perfect score: 54

Sequence: 1 VFRSLKVAESD 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 57534646 residues

Total number of hits satisfying chosen parameters: 7754

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : UniProt-02:*

- 1: uniprot_sprot:*
- 2: uniprot_trembl:*

Pre. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	24	44.4	15 2	QTMIX0
2	23	42.6	15 2	QTMIX0
3	22	40.7	11 2	Q8N103
4	22	40.7	14 2	Q8S2V1
5	22	40.7	15 2	Q8S2V2
6	20	37.0	10 2	Q71VB8
7	20	37.0	10 2	QAB2538
8	19	35.2	8 2	Q9GD00
9	19	35.2	10 2	QTM2ZB
10	19	35.2	10 2	Q9F5W1
11	19	35.2	11 2	P77404
12	19	35.2	13 2	QTRAS8
13	19	35.2	13 2	Q33417
14	19	35.2	13 2	Q9MK6
15	19	35.2	15 2	Q99386
16	19	35.2	15 2	P25451
17	18	33.3	6 2	P7GEF5
18	18	33.3	8 2	Q36898
19	18	33.3	9 2	Q7MV3
20	18	33.3	11 2	Q32704
21	18	33.3	11 2	Q6LDR1
22	18	33.3	11 2	A0A99392
23	18	33.3	11 2	A0A99392
24	18	33.3	14 2	Q71U11
25	18	33.3	14 2	QAC35347
26	18	33.3	14 2	QAC35345
27	18	33.3	15 1	MILT_ONCKE
28	18	33.3	15 2	QTM4Z6
29	18	33.3	15 2	Q9BZNO
30	18	33.3	15 2	Q7r923
31	18	33.3	15 2	Q7rbw7

ALIGNMENTS									
[1]									
RESULT 1									
QTMIX0	ID	QTMIX0	PRELIMINARY;	PRT;	15 AA.				
AC	QTMIX0;								
DT	01-MAR-2004	(TREMBLrel.	26, Created)						
DT	01-MAR-2004	(TREMBLrel.	26, Last sequence update)						
DT	01-MAR-2004	(TREMBLrel.	26, Last annotation update)						
DE	Glycine cleavage system protein H1 and H2 (Fragment).								
DS	Arabidopsis thaliana (Mouse-ear cress).								
OC	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.								
OC	euroids II; Brassicales; Brassicaceae; Arabidopsis.								
OX	NCBI_TaxID=3702;								
RN	NCBI_SEQUENCE.								
RA	Kamo M., Kawakami T., Miyatake N., Tsugita A.;								
RL	Submitted (JUL-1994) to the PIR data bank.								
DR									
FT									
FT	NON_TER	1	1						
FT	NON_TER	15 AA;	15	MW;	5450DFA18A83898C	CRC64;			
FT	SEQUENCE	15 AA;	15						
Query	Match								
		44.4%;	Score 24;	DB 2;	Length 15;				
		62.5%;	Pred. No. 9e+02;	0; Mismatches	3; Indels	0;	Gaps	0;	
Best	Local Similarity								
Matches	5;	Conservative							
QY	1 VFRSLKVA 8								
Db	2								
Db	3 VLEGKVA 10								
[1]									
RESULT 2									
Q86128	ID	Q86128	PRELIMINARY;	PRT;	15 AA.				
AC	Q86128;								
DT	01-NOV-1996	(TREMBLrel.	01, Created)						
DT	01-NOV-1996	(TREMBLrel.	01, Last sequence update)						
DT	01-NOV-1996	(TREMBLrel.	01, Last annotation update)						
DE	N protein (Fragment).								
DS	Vesicular stomatitis virus.								
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales; Rhabdoviridae; Vesiculovirinae.								
OC	NCBI_SEQUENCE.								
RN	NCBI_TaxID=11276;								
RR	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE-80001959; PubMed=89911;								
RA	McGroch D.J.;								
RT	Structure of the gene N: gene NS intercistronic junction in the genome of vesicular stomatitis virus.;								
RT	genome of vesicular stomatitis virus.;								

Q9trlo canis famili
Q9ngs mus sp. maj
Q9gef3 nicotiana g
Q9gef4 nicotiana t
Q97016 nicotiana a
Q9Xgm5 arabidopsis
Q798n7 streptococc
Cab0129 streptoco
P82386 litoria ran
P82387 litoria ran
P82560 streptococc
P82808 rattus norv
Q8wk21 bolidomomas
Q98bm9 raphanus sa
Q53541 bacillus sp
Q61cw3 mus musculu
Q88954 vaccinia vi
Q9ibm5 human immun
Aaa85677 mus muscu

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd:

Om protein - protein search, using sw model
 Run on: January 27, 2005, 19:22:45 ; Search time 39 Seconds
 (without alignments)
 27.138 Million cell updates/sec

Title: US-09-830-837a-77

Perfect score: 54

Sequence: 1 VFRSLKVAESD 11

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext: 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 50 summaries

Database : PIR 79.4

1: pix1: *
 2: pix2: *
 3: pix3: *
 4: pix4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	24	44.4	15	2 PA0036	glycine cleavage system protein H1 and H2 - <i>Arabidopsis thaliana</i> (mouse-ear cress)
2	19	35.2	10	2 JC1367	thyroloberin poten
3	19	35.2	11	2 A26930	ermG leader peptid
4	18	33.3	9	2 PS0253	glycine cleavage s
5	18	33.3	10	2 PT0322	ig heavy chain CRD
6	18	33.3	11	2 PC2372	58K heat shock pro
7	18	33.3	11	2 B29806	acidic proline-ric
8	18	33.3	15	2 PA0093	ematin synthetas
9	18	33.3	15	2 B45133	casein kinase II (
10	17	31.5	11	2 154193	Rheubar blood group
11	17	31.5	12	2 PH1611	Ig H chain V-D-J
12	17	31.5	15	2 JN0730	hypothetical 1.7K
13	16	29.6	7	2 S11797	glucose isomerase
14	16	29.6	8	2 S45651	probable Na+-trans
15	16	29.6	8	2 S37141	rpsA protein - Erw
16	16	29.6	9	2 A44873	caldesmon - rabbit
17	16	29.6	9	2 PT0324	IG heavy chain CRD
18	16	29.6	9	2 S02384	probable membrane
19	16	29.6	9	2 PC0703	ubiquinol-cytochrome
20	16	29.6	10	2 S77990	cytochrome-c oxida
21	16	29.6	11	2 PH0914	T-cell receptor be
22	16	29.6	12	2 PH0936	T-cell receptor be
23	16	29.6	13	2 PT0290	ing heavy chain CRD
24	16	29.6	13	2 H56046	urinary tract stc
25	16	29.6	14	1 LFECRS	pept operon Leade
26	16	29.6	14	2 F90931	pept operon Leade
27	16	29.6	14	2 B85780	pept operon Leade
28	16	29.6	14	2 164815	carbonyl anhydride
29	16	29.6	14	2 AG0705	phenylalanyl-tRNA

Minimum DB seq length: 0

Maximum DB seq length: 15

ALIGNMENTS

RESULT 1

PA0036

glycine cleavage system protein H1 and H2 - *Arabidopsis thaliana* (fragment)

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C;Accession: PA0036

R;Kano, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.

submitted to JPIID, July 1994

A;Description: Separation and characterization of *Arabidopsis* proteins by two-dimension

A;Reference number: PA0001

A;Accession: PA0036

A;Molecule type: protein

A;Residues: 1-15 <RAW>

A;Cross-references: UNIPROT:Q7MIX0

A;Experimental source: stem

Query Match 1 VFRSLKVA 8

Db 3 VLEGKVA 10

RESULT 2

JC1367

thyroloberin potentiating neuropeptide - bovine

N;Alternate names: thyrotropin-releasing hormone potentiating peptide; TRH-potentiating

C;Species: *Bos primigenius taurus* (cattle)

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C;Accession: JC1367

R;Bulant, M.; Ladram, A.; Montagne, J.J.; Delfour, A.; Nicolas, P.

Biochem. Biophys. Res. Commun. 189, 1110-1118, 1992

A;Title: Isolation and amino acid sequence of the TRH-potentiating peptide from bovine

A;Reference number: JC1367; MUID:93111999; PMID:1472021

A;Accession: JC1367

A;Molecule type: protein

A;Residues: 1-10 <RAW>

A;Cross-references: UNIPROT:Q7M228

A;Experimental source: hypothalamus

C;Comment: This neuropeptide corresponds to a region of the rat thyroloberin precursor

C;Function:

A;Description: potentiates thyroloberin-induced thyrotropin secretion

C;Supfamily: thyroloberin

C;Keywords: hypothalamus; neuropeptide

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

August 17, 2005, 15:36:20 ; Search time 169 Seconds

Perfect score: 5537

Sequence: 1 GKKHLGLRLRKKSPEKAPCP... PRVTKRQDLMQQVHPPKTPSV 1036

Scoring table: BLOSUM62

Gapop 10 0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0*, Maximum Match 100*

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgna_6/ptodata/2/pubpaas/us07_PUBCOMB_pep:*

2: /cgna_6/ptodata/2/pubpaas/US07_PUBCOMB_pep:*

3: /cgna_6/ptodata/2/pubpaas/us06_PUBCOMB_pep:*

4: /cgna_6/ptodata/2/pubpaas/us07_PUBCOMB_pep:*

5: /cgna_6/ptodata/2/pubpaas/US07_PUBCOMB_pep:*

6: /cgna_6/ptodata/2/pubpaas/us08_PUBCOMB_pep:*

7: /cgna_6/ptodata/2/pubpaas/us09_PUBCOMB_pep:*

8: /cgna_6/ptodata/2/pubpaas/us09_PUBCOMB_pep:*

9: /cgna_6/ptodata/2/pubpaas/us09_PUBCOMB_pep:*

10: /cgna_6/ptodata/2/pubpaas/us09_PUBCOMB_pep:*

11: /cgna_6/ptodata/2/pubpaas/us09_PUBCOMB_pep:*

12: /cgna_6/ptodata/2/pubpaas/us09_PUBCOMB_pep:*

13: /cgna_6/ptodata/2/pubpaas/us10_PUBCOMB_pep:*

14: /cgna_6/ptodata/2/pubpaas/us10_PUBCOMB_pep:*

15: /cgna_6/ptodata/2/pubpaas/us10_PUBCOMB_pep:*

16: /cgna_6/ptodata/2/pubpaas/us10_PUBCOMB_pep:*

17: /cgna_6/ptodata/2/pubpaas/us10_PUBCOMB_pep:*

18: /cgna_6/ptodata/2/pubpaas/us10_PUBCOMB_pep:*

19: /cgna_6/ptodata/2/pubpaas/us11_PUBCOMB_pep:*

20: /cgna_6/ptodata/2/pubpaas/us11_PUBCOMB_pep:*

21: /cgna_6/ptodata/2/pubpaas/us60_PUBCOMB_pep:*

22: /cgna_6/ptodata/2/pubpaas/us60_PUBCOMB_pep:*

ALIGNMENTS

RESULT 1
US-09-050-854B-7
; Sequence 7, Application US/09060854B

PATENT NO. US20020081703A1
GENERAL INFORMATION:
APPLICANT: ESTELL, David Aaron
TITLE OF INVENTION: Human Protease and Use of Such Protease for Pharmaceutical Applications and for Reducing the Allergenicity of No. US20020081703A1
TITLE OF INVENTION: Applications and for Reducing the Allergenicity of No. US20020081703A1
TITLE OF INVENTION: Proteins

FILE REFERENCE: GC532
CURRENT APPLICATION NUMBER: US/09/060,854B
CURRENT FILING DATE: 1998-04-15
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PastSeq for Windows Version 3.0

SEQ ID NO: 7
LENGTH: 1052
TYPE: PRT
ORGANISM: Homo sapien

US-09-060-854B-7
Query Match: 100.0%; Score: 5537; DB: 9; Length: 1052;
Best Local Similarity: 100.0%; Pred. No: 0;
Matches: 1036; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

Qy 1 GKKHLGLRLRKKSPEKAPCPGCSHILTKVERSSTVVEYVAFNGYFTAKARNPSISSA 60
Db 17 GKKHLGLRLRKKSPEKAPCPGCSHILTKVERSSTVVEYVAFNGYFTAKARNPSISSA 76
Qy 61 LKSSENDWMLPRNPNPSDPSDFEVQIKEROKAGLTLEDHNKIRVTPORKYFRL 120
77 LKSSENDWMLPRNPNPSDPSDFEVQIKEROKAGLTLEDHNKIRVTPORKYFRL 136
Qy 121 KYAESDPPTVNCETRWSQKIQSSRLRASLISLGGFWHATGRHSRRRLRAIPQWAT 180

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5537	100.0	1052	9	US-09-060-854B-7 Sequence 7, Appli
2	5537	100.0	1052	9	US-09-891-711-4 Sequence 4, Appli
3	5537	100.0	1052	15	US-10-665-689-6 Sequence 5, Appli
4	5537	100.0	1052	16	US-10-829-118-4 Sequence 4, Appli
5	5537	100.0	1052	18	US-10-924-92A-208 Sequence 208, Appli
6	5537	100.0	1052	20	US-11-028-208 Sequence 208, Appli
7	5408	97.7	1052	9	US-09-891-711-6 Sequence 6, Appli
8	5408	97.7	1052	16	US-10-829-118-6 Sequence 6, Appli
9	3549	64.1	666	15	US-10-264-2531 Sequence 251, Appli
10	2428	43.9	497	15	US-10-104-047-2347 Sequence 2347, Appli
11	1768.5	31.9	663	15	US-10-424-599-255271 Sequence 255271,

Copyright (c) 1993 - 2005 Compugen Ltd. GenCore version 5.1.6

RESULT 1	US-09-255-502-7
Sequence 7, Application US/09255502	
Patient No. 6218165	
GENERAL INFORMATION:	
APPLICANT: Estell, David	
APPLICANT: Harding, Fiona	
TITLE OF INVENTION: Mutant Proteins Having Lower Allergenic Responses in Humans and Methods for Constructing, Identifying and Producing Such Proteins	
FILE REFERENCER: GC 527-D2	
CURRENT APPLICATION NUMBER: US/09/255,502	
CURRENT FILING DATE: 1999-02-23	
PRIOR APPLICATION NUMBER: 09/050,872	
PRIOR FILING DATE: 1998-04-15	
NUMBER OF SBO ID NOS: 7	
SOFTWARE: Patentin Ver. 2.0	
SBO ID NO 7	
LENGTH: 1052	
TYPE: PRT	
ORGANISM: Homo sapiens	
US-09-255-502-7	
Query Match 100.0%; Score 5537; DB 3; Length 1052;	
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;	
Matches 1036; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy 1 GKKHLGDRLEKSPKACPCGCSHLTKVPSSTVYETVAFNQYPTAKARNPISSA 60	Db 17 GKKHLGDRLEKSPKACPCGCSHLTKVRSSTVYETVAFNQYPTAKARNPISSA 76
Qy 61 LKSEVDNWRITPRNPPSSQYPSDSEVIIQKERKQAGLTLIEDHNPKRVTQPKVFRSL 120	Db 77 LKSEVDNWRITPRNPPSSQYPSDSEVIIQKERKQAGLTLIEDHNPKRVTQPKVFRSL 136
Qy 121 KYAESDPTVPRNETWSOKQSSRFLRRASLSSLGGFWHATGRSSRLRAIPQVACT 180	Db 137 KYAESDPTVPRNETWSOKQSSRFLRRASLSSLGGFWHATGRSSRLRAIPQVACT 196
Qy 181 LOADVILWOMSYTGANRVAVFDGLSERKHPKVNKTWTNTLDDGLGHGTPVAGY 240	Db 197 LOADVILWOMGTGANRVAVFDGLSERKHPKVNKTWTNTLDDGLGHGTPVAGY 256
Qy 241 IASMRBCQGAPDAELHIFKVFTNNQVSYTTSWFLDAFNAYAIIKKCDVNLNISIGGDFMDH 300	Db 317 IASMRBCQGAPDAELHIFKVFTNNQVSYTTSWFLDAFNAYAIIKKCDVNLNISIGGDFMDH 316
Db 301 PFDVKWMLTANNVIVMSAIGNDGLYLTNNPADQMDVYGGGGIDPFDNIARESRGMT 360	Db 317 PFDVKWMLTANNVIVMSAIGNDGLYLTNNPADQMDVYGGGGIDPFDNIARESRGMT 376

Scoring table:						
Searched:						GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.						Copyright (c) 1993 - 2005 Compugen Ltd.
OM protein - protein search, using sw model						OM protein - protein search, using sw model
Run on:						Run on:
August 17, 2005, 15:31:04 ; Search time 183 Seconds						August 17, 2005, 15:31:04 ; Search time 183 Seconds
Post-processing: (without alignments)						Post-processing: (without alignments)
Database : Maximum Match 0%						Database : Maximum Match 0%
UniProt 03:*						UniProt 03:*
1: uniprot_sprot:*						1: uniprot_sprot:*
2: uniprot_trembl:*						2: uniprot_trembl:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.						
SUMMARIES						
Result No.	Score	Query Match	Length	DB ID	Description	
1	5537	100.0	1052	1	MS1P_HUMAN	
2	5408	97.7	1052	1	MS1P_CRIGR	
3	5396	97.5	1052	1	MS1P_RAT	
4	5360	96.8	1052	1	MS1P_MOUSE	
5	5357	96.7	1052	2	06PG67	
6	5339	96.4	2	08QUB0	MS1P_MOUSE	
7	4770	86.1	1074	2	06XJU1	
8	4522	81.7	950	2	06AX98	
9	2835	51.2	2	08TAN4	MS1P_HUMAN	
10	2815.5	50.8	1238	2	07PVBO	
11	2377.5	42.9	1012	2	09VP10	
12	2269	41.0	2288	2	08L7B7	
13	355	6.4	1239	2	Q9FBZ4	
14	346.5	6.3	1722	2	09NDZ2	
15	339.5	6.1	381	2	Q84fP8	
16	338.5	6.1	381	1	SUBN_BACNA	
17	338.5	6.1	381	1	SUBT_BACSA	
18	338.5	6.1	381	1	SUBT_BACST	
19	337.5	6.1	381	1	SUBT_BACSU	
20	337.5	6.1	1139	2	082119	
21	336	6.1	1102	2	P95684	
22	335.5	6.1	1245	2	Q9RL54	
23	333.5	6.0	379	2	Q9Rdf4	
24	333.5	6.0	379	2	Q9Rdf4	
25	333.5	6.0	379	2	Q9Rdf4	
26	329	5.9	379	1	SUBT_BACLI	
27	327.5	5.9	627	2	Q9Rdf4	
28	322	5.8	382	2	Q45522	
29	321	5.8	379	2	Q45301	
30	321	5.8	379	2	Q6Cn9	
31	321	5.8	379	2	Q6PNN5	

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 17, 2005, 15:19:31 ; Search time 48 seconds
 (without alignments)
 2076.678 Million cell updates/sec

Title: US-09-830-837a-6_COPY_17_1052

Perfect score: 5537

Sequence: 1 GKKHGLDRLEKKSFPEKRPCP.....PRVKRFPQLMQQVHPPKPSV 1036

Scoring table: BLOSUM62

Gapop 10.0 , Gapext: 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : PIR 791*

1: piri: *
 2: pir2: *
 3: pir3: *
 4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	5408	97.7	1052	2	T17093 intraluminal subtilisin-like proteinase SIP, membrane-bound - Chinese hamster
2	1716	31.0	318	2	T43492 hypothetical protease (EC 3.4.2.1)
3	338.5	6.1	381	1	SUBSS subtilisin (EC 3.4.2.1)
4	338.5	6.1	381	2	JQ1487 subtilisin (EC 3.4.2.1)
5	338.5	6.1	381	2	JH0778 subtilisin (EC 3.4.2.1)
6	337.5	6.1	381	1	SUBSI subtilisin (EC 3.4.2.1)
7	329	5.9	379	1	SUBSI subtilisin (EC 3.4.2.1)
8	327.5	5.9	627	2	D75393 serine protease, subtilisin (EC 3.4.2.1)
9	322	5.8	382	2	I39780 subtilisin (EC 3.4.2.1)
10	318.5	5.8	382	1	SUBSN subtilisin (EC 3.4.2.1)
11	312.5	5.6	272	2	A23624 subtilisin (EC 3.4.2.1)
12	308	5.6	274	1	SUBSD subtilisin (EC 3.4.2.1)
13	304.5	5.5	323	2	I39867 microbial serine protease, subtilisin (EC 3.4.2.1)
14	299	5.4	326	1	C41335 subtilisin (EC 3.4.2.1)
15	295	5.3	275	2	JCL1085 subtilisin (EC 3.4.2.1)
16	294	5.3	321	2	JCG460 intracellular alkaline proteinase, subtilisin (EC 3.4.2.1)
17	293.5	5.3	322	2	G83922 microbial serine protease, subtilisin (EC 3.4.2.1)
18	289	5.2	319	2	I39866 alkaline proteins, subtilisin (EC 3.4.2.1)
19	285.5	5.2	374	2	I39781 high-alkaline serine protease, subtilisin-like proteinase, subtilase family protein, cysteine-dependent
20	284.5	5.1	321	1	S27501 A49778 pyrolysin (EC 3.4.2.1)
21	283	5.1	380	2	A49778 T28159 intracellular alkaline proteinase, subtilisin (EC 3.4.2.1)
22	281.5	5.1	1398	2	T28159 alkaline proteinase, subtilisin (EC 3.4.2.1)
23	271	4.9	384	2	JCG4802 probable protease, high-alkaline serine protease, subtilisin-like proteinase, subtilase family protein, cysteine-dependent
24	270.5	4.9	328	2	G95392 T28159 probable protease, high-alkaline serine protease, subtilisin-like proteinase, subtilase family protein, cysteine-dependent
25	266	4.8	378	2	A33973 S44131 A12007 C84120 probable protease, high-alkaline serine protease, subtilisin-like proteinase, subtilase family protein, cysteine-dependent
26	260.5	4.7	682	2	S44131 A12007 C84120 probable protease, high-alkaline serine protease, subtilisin-like proteinase, subtilase family protein, cysteine-dependent
27	257	4.6	1448	2	A12007 C84120 probable protease, high-alkaline serine protease, subtilisin-like proteinase, subtilase family protein, cysteine-dependent
28	256	4.6	757	2	C84120 probable protease, high-alkaline serine protease, subtilisin-like proteinase, subtilase family protein, cysteine-dependent
29	253	4.6	397	2	JW0075 probable protease, high-alkaline serine protease, subtilisin-like proteinase, subtilase family protein, cysteine-dependent

RESULT 1
 T17093
 intraluminal subtilisin-like proteinase SIP, membrane-bound - Chinese hamster
 C;Species: *Cricetulus griseus* (Chinese hamster)
 C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
 C;Accession: T17093
 R;Sakai, J.; Rawson, R.B.; Espenashade, P.J.; Cheng, D.; Seegmiller, A.C.; Goldstein, J.
 Molecular Cell 2, 505-514, 1998
 A;Title: Molecular identification of the sterol-regulated luminal protease that cleaves
 A;Reference number: 218677; MUID:99026600; PMID:9809072
 A;Accession: T17093
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-1052 <SAK>
 A;Cross-references: UNIPROT:Q9Z2A8; EMBL:AF078105; NID:93892203; PID:93892204; PIDN: AAC
 A;Experimental source: strain 25-RA
 C;Function:
 A;Description: SIP is proposed to be a sterol-regulated protease that controls lipid me
 A;Note: SIP cleaves sterol regulatory element binding proteins (SREBPs) in the ER lumen

Query Match 97.7%; Pred. No. 0; Score 5408; DB 2; Length 1052;
 Best Local Similarity 97.1%;保守性 15; Mi Matches 15; Indels 0; Gaps 0;
 Matches 1006; Conservative 15; Mi Matches 15; Indels 0; Gaps 0;

QY 1 GKKHGLDRLEKKSFPEKRPCP...
 Db 17 GKKHGLDRLEKKSFPEKRPCP...
 QY 61 LKSSAVDNRWIPRANPSSD...
 Db 77 LKSSAVDNRWIPRANPSSD...
 QY 121 KYAESDPTVPCNETRWSK...
 Db 137 KFAESDPTVPCNETRWSK...
 QY 181 LQADTYWQMYTG...
 Db 191 LQADTYWQMYTG...
 QY 240 181 LQADTYWQMYTG...
 Db 196 181 LQADTYWQMYTG...
 Db 197 LQADTYWQMYTG...
 QY 241 IASMECGQFAPDAELH...
 Db 257 IASMECGQFAPDAELH...
 QY 301 PFDVKWMLTANNYTM...
 Db 317 PFDVKWMLTANNYTM...
 QY 361 TWELGGYSGM...
 Db 377 TWELGGYSGM...
 QY 420 TWELGGYSGM...
 Db 436 TWELGGYSGM...
 QY 436 TWELGGYSGM...

Copyright (c) 1993 - 2005 Compugen Ltd.

Om protein - protein search, using sw model

Run on:

August 17, 2005, 15:18:18 ; Search time 169 Seconds

(Without alignments)

2370.910 Million cell updates/sec

Title: US-09-830-837A-6_COPY_17_1052

Perfect score: 5537

Sequence: GKKHLGRLEKKSFEKQCP.....PRVKRPQLMQQVHPPKTPSV 1036

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseq1980s:*
- 2: geneseq2000s:*
- 3: geneseq2016:*
- 4: geneseq2002a:*
- 5: geneseq2003a:*
- 6: geneseq2003b:*
- 7: geneseq2004b:*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description	RESULT
1	5537	100.0	1052	AAB06334	1 AAB06334
2	5537	100.0	1052	3 AAY4619	1 AAB06334 standard; protein; 1052 AA.
3	5537	100.0	1052	3 AAY66982	1 XX
4	5537	100.0	1052	3 AAY84227	1 AC AAB06334;
5	5537	100.0	1052	4 AAU8520	1 XX
6	5537	100.0	1052	5 AAE14527	1 DT 03-OCT-2000 (first entry)
7	5537	100.0	1052	5 ABG1096	1 XX
8	5408	97.7	1052	3 AAY84228	1 XX
9	5408	97.7	1052	5 AAE14528	1 XX
10	5396	97.5	1052	3 AAB06335	1 PA (RECL-) INST RECH CLINIQUES MONTREAL.
11	5360	96.8	1052	3 AAB06336	1 XX
12	3549	64.1	666	5 ABP0255	1 PI Se-dah N, Chretien M, Marcinkiewicz M, Laakonen R, Davignon J, XX
13	2428	43.9	497	7 ADB64193	1 DR WPI: 2000-365601/31.
14	2368.5	42.7	992	4 AAB20015	1 DR N-PSDB; AAA57197.
15	2231	40.3	492	4 AAB3047	1 DR
16	606	10.9	129	5 ADK34880	1 DR
17	585	10.6	125	5 ADK36735	1 DR
18	463	8.4	126	4 AAO1619	1 DR
19	343.5	6.2	392	6 ABP9571	1 PT Novel soluble proteic fragment of subtilisin-kexin isoenzyme for producing a polypeptide useful for treating hypercholesterolemia, liver
20	340.5	6.1	381	1 AAB60571	1 PT
21	340.5	6.1	381	1 AAB70744	1 PT
22	340.5	6.1	381	2 AAR07970	1 PT
23	338.5	6.1	352	2 AAW46600	1 PT
24	338.5	6.1	387	2 ABP9567	1 PT
25	338.5	6.1	388	6 ABP9569	1 PT

ALIGNMENTS

Abp59565	Protein a
Aaw6595	Amino aci
Aap71060	Wild-type
Aar0737	Subtilisi
Aar24131	Bacterial
Aar3463	Bacillus
Aar74224	B. subtili
Aay79229	Bacillus
Aay70053	Bacillus
Aab81180	Transglut
Abu07391	Foreign P
Aar51928	Bacillus
Aae29944	Bacillus
Aar19063	Bacillus
Aar19062	Bacillus
Aar52643	Bacillus
Aar51924	Bacillus
Aar51929	Bacillus
Aar1920	Bacillus
Aar51927	Bacillus